



GC830-PCT.txt
SEQUENCE LISTING

<110> Genencor International, Inc.
Dunn-Coleman, N.
Shetty, J.
Duan, G.
Sung, A.
Qian, Y.

<120> A Method for the Preparation of a High Purity Rice Protein Concentrate

<130> GC830-PCT

<140> PCT/US05/05320

<141> 2005-02-17

<150> US 60/547,153

<151> 2004-02-23

<160> 11

<170> PatentIn version 3.2

<210> 1

<211> 2103

<212> DNA

<213> Humicola grisea var. thermoidea

<400> 1

atgcatacct	tctccaagct	cctcgctcctg	ggctctgccc	tccagttctgc	cctcggggcgg	60
cctcacggct	cttcgcgtct	ccaggaacgc	gctgccgttg	ataccttcat	caacaccgag	120
aagcccatcg	catggaacaa	gctgctcgcc	aacatcggcc	ctaaccggcaa	agccgctccc	180
ggtgccgcgcg	ccggcggttg	gattgccagc	ccttccagga	cggaccctcc	ttgtacgtgg	240
tggcatggaa	tggacccaag	agactggttt	tagatgaaag	agagtttctg	ctaaccgcca	300
cacccagact	tcttcacctg	gacccgcgat	gccgccctgg	tcctcaccgg	catcatcgag	360
tcccttggcc	acaactacaa	caccaccctg	cagaccgtca	tccagaacta	cgtcgcgtcg	420
caggccaagc	tgcagcaggt	ctcgaacccc	tcgggaacct	tcgccgacgg	ctcgggtctc	480
ggtgaggcca	agttcaatgt	cgacctcact	gccttcaactg	gcgaatgggg	tcgccctcag	540
agggacggcc	cgcccctgcg	cgccatcgct	ctcatccagt	acgccaaagt	gctgatcgcc	600
aacggctaca	agagcacggc	caagagcgtc	gtctggcccc	tcgtcaagaa	cgatctcgcc	660
tacacggccc	agtactggaa	cgagaccggc	ttcgatctct	gggaggaggt	ccccggcagc	720
tcgtttctta	ccatcgccag	ctctcacagg	ggtgagtcac	ttattgttca	gtgttttctc	780
attgaataat	taccggaatg	ccactgacgc	caaacagctc	tgactgaggg	tgcttacctc	840
gccgctcagc	tcgacaccga	gtgccgcgcc	tgcacgaccg	tcgcccctca	ggttctgtgc	900
ttccagcagg	ccttctggaa	ctccaagggc	aactatgtcg	tctccaacag	taagatccct	960
acaccaacaa	aaaaaatcga	aaaggaacgt	tagctgaccc	ttctagtcaa	cggcggcgag	1020
tatcgctccg	gcaaggacgc	caactcgatc	ctggcgctca	tccacaactt	cgaccctgag	1080
gccggctgcg	acaacctgac	cttccagccc	tgcagcgagc	gcgccctggc	caaccacaag	1140
gcctatgtcg	actcgttccg	caacctctac	gccatcaaca	agggcatcgc	ccagggaag	1200
gccgttgccg	tcggccgcta	ctcggaggat	gtctactaca	acggcaaccc	gtggtacctg	1260
gccaactttg	ccgccgcga	gcagctctac	gacgccatct	acgtgtggaa	caagcagggc	1320
tccatcaccg	tgacctcggt	ctccctgccc	ttcttccgcg	accttgtctc	gtcggtcagc	1380
accggcacct	actccaagag	cagctcgacc	ttcaccaaca	tcgtcaacgc	cgtcaaggcc	1440
tacgccgacg	gcttcatcga	ggtggcggcc	aagtacaccc	cgtccaacgg	cgcgctcgcc	1500
gagcagtagc	accgcaacac	gggcaagccc	gactcgccgc	ccgacctgac	gtggtcgtac	1560
tcggccttcc	tctcgcccat	cgaccgcgcg	gcgggtctcg	tcccccgag	ctggcggggc	1620
agcgtggcca	agagccagct	gccgtccacc	tgtctcgcca	tcgaggtcgc	cggcacctac	1680
gtcgccgcca	cgagcacctc	gttcccgtcc	aagcagaccc	cgaaccctc	cgcggcgccc	1740

GC830-PCT.txt

```

tccccgtccc cctacccgac cgcctgcgcg gacgctagcg aggtgtacgt caccttcaac 1800
gagcgcgtgt cgaccgcgtg gggcgagacc atcaaggtagg tgggcaacgt gccggcgctg 1860
gggaactggg acacgtccaa ggcggtgacc ctgtcggcca gcgggtacaa gtcgaatgat 1920
ccccctctgga gcatcacggt gcccatacaag gcgacgggct cggccgtgca gtacaagtat 1980
atcaaggtagg gcaccaacgg gaagattact tgggagtcgg accccaacag gagcattacc 2040
ctgcagacgg cgtcgtctgc gggcaagtgc gccgcgcaga cggatgaatga ttcgtggcgt 2100
taa 2103

```

<210> 2

<211> 634

<212> PRT

<213> Humicola grisea var. thermoidea

<400> 2

```

Met His Thr Phe Ser Lys Leu Leu Val Leu Gly Ser Ala Val Gln Ser
1      5      10      15
Ala Leu Gly Arg Pro His Gly Ser Ser Arg Leu Gln Glu Arg Ala Ala
20      25      30
Val Asp Thr Phe Ile Asn Thr Glu Lys Pro Ile Ala Trp Asn Lys Leu
35      40      45
Leu Ala Asn Ile Gly Pro Asn Gly Lys Ala Ala Pro Gly Ala Ala Ala
50      55      60
Gly Val Val Ile Ala Ser Pro Ser Arg Thr Asp Pro Pro Tyr Phe Phe
65      70      75      80
Thr Trp Thr Arg Asp Ala Ala Leu Val Leu Thr Gly Ile Ile Glu Ser
85      90      95
Leu Gly His Asn Tyr Asn Thr Thr Leu Gln Thr Val Ile Gln Asn Tyr
100     105     110
Val Ala Ser Gln Ala Lys Leu Gln Val Ser Asn Pro Ser Gly Thr
115     120     125
Phe Ala Asp Gly Ser Gly Leu Gly Glu Ala Lys Phe Asn Val Asp Leu
130     135     140
Thr Ala Phe Thr Gly Glu Trp Gly Arg Pro Gln Arg Asp Gly Pro Pro
145     150     155     160
Leu Arg Ala Ile Ala Leu Ile Gln Tyr Ala Lys Trp Leu Ile Ala Asn
165     170     175
Gly Tyr Lys Ser Thr Ala Lys Ser Val Val Trp Pro Val Val Lys Asn
180     185     190
Asp Leu Ala Tyr Thr Ala Gln Tyr Trp Asn Glu Thr Gly Phe Asp Leu
195     200     205
Trp Glu Glu Val Pro Gly Ser Ser Phe Phe Thr Ile Ala Ser Ser His
210     215     220
Arg Ala Leu Thr Glu Gly Ala Tyr Leu Ala Ala Gln Leu Asp Thr Glu
225     230     235     240
Cys Arg Ala Cys Thr Thr Val Ala Pro Gln Val Leu Cys Phe Gln Gln
245     250     255
Ala Phe Trp Asn Ser Lys Gly Asn Tyr Val Val Ser Asn Ile Asn Gly
260     265     270
Gly Glu Tyr Arg Ser Gly Lys Asp Ala Asn Ser Ile Leu Ala Ser Ile
275     280     285
His Asn Phe Asp Pro Glu Ala Gly Cys Asp Asn Leu Thr Phe Gln Pro
290     295     300
Cys Ser Glu Arg Ala Leu Ala Asn His Lys Ala Tyr Val Asp Ser Phe
305     310     315     320
Arg Asn Leu Tyr Ala Ile Asn Lys Gly Ile Ala Gln Gly Lys Ala Val
325     330     335
Ala Val Gly Arg Tyr Ser Glu Asp Val Tyr Tyr Asn Gly Asn Pro Trp
340     345     350
Tyr Leu Ala Asn Phe Ala Ala Ala Glu Gln Leu Tyr Asp Ala Ile Tyr

```

GC830-PCT.txt

```

      355      360      365
Val Trp Asn Lys Gln Gly Ser Ile Thr Val Thr Ser Val Ser Leu Pro
370      375      380
Phe Phe Arg Asp Leu Val Ser Ser Val Ser Thr Gly Thr Tyr Ser Lys
385      390      395      400
Ser Ser Ser Thr Phe Thr Asn Ile Val Asn Ala Val Lys Ala Tyr Ala
405      410      415
Asp Gly Phe Ile Glu Val Ala Ala Lys Tyr Thr Pro Ser Asn Gly Ala
420      425      430
Leu Ala Glu Gln Tyr Asp Arg Asn Thr Gly Lys Pro Asp Ser Ala Ala
435      440      445
Asp Leu Thr Trp Ser Tyr Ser Ala Phe Leu Ser Ala Ile Asp Arg Arg
450      455      460
Ala Gly Leu Val Pro Pro Ser Trp Arg Ala Ser Val Ala Lys Ser Gln
465      470      475      480
Leu Pro Ser Thr Cys Ser Arg Ile Glu Val Ala Gly Thr Tyr Val Ala
485      490      495
Ala Thr Ser Thr Ser Phe Pro Ser Lys Gln Thr Pro Asn Pro Ser Ala
500      505      510
Ala Pro Ser Pro Ser Pro Tyr Pro Thr Ala Cys Ala Asp Ala Ser Glu
515      520      525
Val Tyr Val Thr Phe Asn Glu Arg Val Ser Thr Ala Trp Gly Glu Thr
530      535      540
Ile Lys Val Val Gly Asn Val Pro Ala Leu Gly Asn Trp Asp Thr Ser
545      550      555      560
Lys Ala Val Thr Leu Ser Ala Ser Gly Tyr Lys Ser Asn Asp Pro Leu
565      570      575
Trp Ser Ile Thr Val Pro Ile Lys Ala Thr Gly Ser Ala Val Gln Tyr
580      585      590
Lys Tyr Ile Lys Val Gly Thr Asn Gly Lys Ile Thr Trp Glu Ser Asp
595      600      605
Pro Asn Arg Ser Ile Thr Leu Gln Thr Ala Ser Ser Ala Gly Lys Cys
610      615      620
Ala Ala Gln Thr Val Asn Asp Ser Trp Arg
625      630

```

<210> 3

<211> 604

<212> PRT

<213> Humicola grisea var. thermoidea

<400> 3

```

Ala Ala Val Asp Thr Phe Ile Asn Thr Glu Lys Pro Ile Ala Trp Asn
1      5      10      15
Lys Leu Leu Ala Asn Ile Gly Pro Asn Gly Lys Ala Ala Pro Gly Ala
20      25      30
Ala Ala Gly Val Val Ile Ala Ser Pro Ser Arg Thr Asp Pro Pro Tyr
35      40      45
Phe Phe Thr Trp Thr Arg Asp Ala Ala Leu Val Leu Thr Gly Ile Ile
50      55      60
Glu Ser Leu Gly His Asn Tyr Asn Thr Thr Leu Gln Thr Val Ile Gln
65      70      75      80
Asn Tyr Val Ala Ser Gln Ala Lys Leu Gln Gln Val Ser Asn Pro Ser
85      90      95
Gly Thr Phe Ala Asp Gly Ser Gly Leu Gly Glu Ala Lys Phe Asn Val
100      105      110
Asp Leu Thr Ala Phe Thr Gly Glu Trp Gly Arg Pro Gln Arg Asp Gly
115      120      125
Pro Pro Leu Arg Ala Ile Ala Leu Ile Gln Tyr Ala Lys Trp Leu Ile

```

130						135						140					
Ala 145	Asn	Gly	Tyr	Lys	Ser 150	Thr	Ala	Lys	Ser	Val 155	Val	Trp	Pro	Val	Val 160		
Lys	Asn	Asp	Leu	Ala 165	Tyr	Thr	Ala	Gln	Tyr 170	Trp	Asn	Glu	Thr	Gly 175	Phe		
Asp	Leu	Trp	Glu 180	Glu	Val	Pro	Gly	Ser 185	Ser	Phe	Phe	Thr	Ile 190	Ala	Ser		
Ser	His	Arg 195	Ala	Leu	Thr	Glu	Gly 200	Ala	Tyr	Leu	Ala	Ala 205	Gln	Leu	Asp		
Thr	Glu 210	Cys	Arg	Ala	Cys	Thr 215	Thr	Val	Ala	Pro	Gln 220	Val	Leu	Cys	Phe		
Gln 225	Gln	Ala	Phe	Trp	Asn 230	Ser	Lys	Gly	Asn 235	Tyr	Val	Val	Ser	Asn	Ile 240		
Asn	Gly	Gly	Glu	Tyr 245	Arg	Ser	Gly	Lys	Asp 250	Ala	Asn	Ser	Ile	Leu 255	Ala		
Ser	Ile	His 260	Asn	Phe	Asp	Pro	Glu 265	Ala	Gly	Cys	Asp	Asn	Leu 270	Thr	Phe		
Gln	Pro	Cys 275	Ser	Glu	Arg	Ala 280	Leu	Ala	Asn	His	Lys	Ala 285	Tyr	Val	Asp		
Ser	Phe 290	Arg	Asn	Leu	Tyr	Ala 295	Ile	Asn	Lys	Gly	Ile 300	Ala	Gln	Gly	Lys		
Ala 305	Val	Ala	Val	Gly	Arg 310	Tyr	Ser	Glu	Asp 315	Val	Tyr	Tyr	Asn	Gly	Asn 320		
Pro	Trp	Tyr	Leu	Ala 325	Asn	Phe	Ala	Ala	Ala 330	Glu	Gln	Leu	Tyr	Asp 335	Ala		
Ile	Tyr	Val 340	Trp	Asn	Lys	Gln	Gly	Ser 345	Ile	Thr	Val	Thr	Ser 350	Val	Ser		
Leu	Pro	Phe 355	Phe	Arg	Asp	Leu	Val 360	Ser	Ser	Val	Ser	Thr 365	Gly	Thr	Tyr		
Ser	Lys 370	Ser	Ser	Ser	Thr	Phe 375	Thr	Asn	Ile	Val	Asn 380	Ala	Val	Lys	Ala		
Tyr 385	Ala	Asp	Gly	Phe	Ile 390	Glu	Val	Ala	Ala 395	Lys	Tyr	Thr	Pro	Ser	Asn 400		
Gly	Ala	Leu	Ala	Glu 405	Gln	Tyr	Asp	Arg	Asn 410	Thr	Gly	Lys	Pro	Asp 415	Ser		
Ala	Ala	Asp 420	Leu	Thr	Trp	Ser	Tyr	Ser 425	Ala	Phe	Leu	Ser	Ala 430	Ile	Asp		
Arg	Arg	Ala 435	Gly	Leu	Val	Pro	Pro 440	Ser	Trp	Arg	Ala	Ser 445	Val	Ala	Lys		
Ser	Gln 450	Leu	Pro	Ser	Thr	Cys 455	Ser	Arg	Ile	Glu	Val 460	Ala	Gly	Thr	Tyr		
Val 465	Ala	Ala	Thr	Ser	Thr 470	Ser	Phe	Pro	Ser	Lys 475	Gln	Thr	Pro	Asn	Pro 480		
Ser	Ala	Ala	Pro	Ser 485	Pro	Ser	Pro	Tyr	Pro 490	Thr	Ala	Cys	Ala 495	Asp	Ala		
Ser	Glu	Val 500	Tyr	Val	Thr	Phe	Asn 505	Glu	Arg	Val	Ser	Thr	Ala 510	Trp	Gly		
Glu	Thr 515	Ile	Lys	Val	Val	Gly	Asn 520	Val	Pro	Ala	Leu	Gly 525	Asn	Trp	Asp		
Thr	Ser 530	Lys	Ala	Val	Thr	Leu 535	Ser	Ala	Ser	Gly	Tyr 540	Lys	Ser	Asn	Asp		
Pro 545	Leu	Trp	Ser	Ile	Thr 550	Val	Pro	Ile	Lys	Ala 555	Thr	Gly	Ser	Ala	Val 560		
Gln	Tyr	Lys	Tyr	Ile 565	Lys	Val	Gly	Thr	Asn 570	Gly	Lys	Ile	Thr	Trp 575	Glu		
Ser	Asp	Pro	Asn 580	Arg	Ser	Ile	Thr	Leu 585	Gln	Thr	Ala	Ser	Ser 590	Ala	Gly		
Lys	Cys	Ala 595	Ala	Gln	Thr	Val	Asn 600	Asp	Ser	Trp	Arg						

GC830-PCT.txt

<210> 4
 <211> 10739
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> pTrex3g_N13 plasmid

<400> 4
 aagcttacta gtactttctcg agctctgtac atgtccggtc ggcagctacg cgtatcgatg 60
 gcgccagctg caggcgggccg cctgcagcca cttgcagtc cgtggaattc tcacggtgaa 120
 tgtaggcctt ttgtagggtta ggaattgtca ctcaagcacc cccaacctcc attacgcctc 180
 ccccatagag ttcccaatca gtgagtcatg gcaactgttct caaatagatt ggggagaagt 240
 tgacttccgc ccagagctga aggtcgca accgcatgat atagggtcgg caacggcaaa 300
 aaagcacgtg gctcaccgaa aagcaagatg tttgcgatct aacatccagg aacctggata 360
 catccatcat cagcacgac cactttgatc tgctggtaaa ctctgtattcg ccctaaaccg 420
 aagtgcgtgg taaatctaca cgtgggcccc tttcgggtata ctgcgtgtgt cttctctagg 480
 tgccattctt ttcccttctt ctagtgttga attgtttgtg ttggagtccg agctgtaact 540
 acctctgaat ctctggagaa tgggtggacta acgactaccg tgcacctgca tcatgtatat 600
 aatagtgatc ctgagaaggg ggggtttggag caatgtggga ctttgatggg catcaaacaa 660
 agaacgaaga cgcctctttt gcaaagtttt gtttcggcta cgggtgaagaa ctggatactt 720
 gttgtgtctt ctgtgtattt ttgtggcaac aagaggccag agacaatcta ttcaaacc 780
 aagcttgctc ttttgagcta caagaacctg tgggttatat atctagagtt gtgaagtcgg 840
 taatcccgtc gtatagtaat acgagtcgca tctaaatact ccgaagctgc tgcgaaccgc 900
 gagaatcgag atgtgctgga aagcttctag cgagcggcta aattagcatg aaaggctatg 960
 agaaattctg gagacggctt gttgaatcat ggcgttccat tcttcgacaa gcaaagcgtt 1020
 ccgtcgcagt agcaggcact cattcccga aaaactcgga gattccctaag tagcgatgga 1080
 accggaataa tataatagga aatacattga gttgcctcga cggttgcaat gcaggggtac 1140
 tgagcttgga cataactgtt ccgtacccca cctcttctca acctttggcg tttccctgat 1200
 tcagcgtacc cgtacaagtc gtaataccta ttaaccgaga ctgaccggac gtgttttgcc 1260
 cttcatttgg agaaataatg tcattgcgat gtgtaatttg cctgcttgac cgactggggc 1320
 tgttcgaagc ccgaatgtag gattgttata cgaactctgc tcgtagaggc atgttgtgaa 1380
 tctgtgtcgg gcaggacacg cctcgaaggt tcacggcaag ggaaaccacc gatagcagtg 1440
 tctagtagca acctgtaaag ccgcaatgca gcatcactgg aaaatacaaa ccaatggcta 1500
 aaagtacata agttaatgcc taaagaagtc atataccagc ggctaataat tgtacaatca 1560
 agtggtctaaa cgtaccgtaa tttgccaaac gcttgtgggg ttgcagaagc aacggcaaa 1620
 ccccaacttc ccacgtttgt ttcttctact agtccaatct cagctggtga tcccccaatt 1680
 gggtcgcttg tttgttccgg tgaagtgaag gaagacagag gtaagaatgt ctgactcgga 1740
 gcgtttttgca tacaaccaag ggcagtgatg gaagacagtg aaatgttgac attcaaggag 1800
 tattttagcca gggatgcttg agtgatcgtt agtaggaggt ttgtctgccc atacgacgaa 1860
 tactgtatag tcacttctga tgaagtgttc catattgaaa tgtaaagtgc gcaactgaaca 1920
 ggcaaaagat tgagttgaaa ctgcctaaga tctcggggccc tcgggccttc ggcctttggg 1980
 tgtacatgtt tgtgtcctcg gcaaagtcaa agtgtggtag gatcgaacac actgctgcct 2040
 ttaccaagca gctgagggtg tgtgataggc aaatgttcag gggccactgc atgggttctga 2100
 atagaaagag aagcttagcc aagaacaata gccgataaag atagcctcat taaacggaat 2160
 gagctagtag gcaaagtcag cgaatgtgta tatataaagg ttcgagggtc gtgcctccct 2220
 catgctctcc ccacttactc atcaactcag atcctccagg agacttgtag accatctttt 2280
 gagcacaga aaccacaatg tcaaccatca caagtttgta caaaaaagca ggctccgcgg 2340
 ccgccccctt caacatgcat accttctcca agctcctcgt cctgggctct gccgtccagt 2400
 ctgcctcgg gcggcctcac ggctcttcgc gtctccagga acgcgctgcc gttgatacct 2460
 tcatcaacac cgagaagccc atcgcatgga acaagctgct cgccaacatc ggccctaacg 2520
 gcaaagccgc tcccgggtgc gccgcggcg ttgtgattgc cagcccttc aggacggacc 2580
 ctcttgttac gtggtggcat ggaatggacc caagagactg gttttagatg aaagagagtt 2640
 tctgctaacc gccacacca gacttcttca cctggaccgc cgatgccgcc ctggctccta 2700
 ccggcatcat cgagtcctt ggccacaact acaacaccac cctgcagacc gtcatccaga 2760
 actacgtcgc gtcgcaggcc aagctgcagc aggtctcgaa cccctcggga accttcgccc 2820
 acggctcggg tctcgggtgag gccaaagttca atgtcgacct cactgccttc actggcgaat 2880
 ggggtcggcc tcagagggac ggcccggccc tgcgcgccat cgctctcatc cagtacgcca 2940

GC830-PCT.txt

agtggctgat	cgccaacggc	tacaagagca	cggccaagag	cgctcgtctg	cccgtcgtca	3000
agaacgatct	cgcctacacg	gcccagtaact	ggaacgagac	cggcttcgat	ctctgggagg	3060
aggtccccgg	cagctcgttc	tttaccatcg	ccagctctca	caggggtgag	tcattttattg	3120
ttcagtgttt	tctcattgaa	taattaccgg	aatgccactg	acgccaacaa	gctctgactg	3180
aggggtgctta	cctcgccgct	cagctcgaca	ccgagtgcg	cgcctgcacg	accgtcgccc	3240
ctcaggttct	gtgcttccag	caggccttct	ggaactccaa	gggcaactat	gtcgtctcca	3300
acagtaagat	ccctacacca	acaaaaaaaa	tcgaaaagga	acgttagctg	acccttctag	3360
tcaacggcgg	cgagtatcgc	tccggcaagg	acgccaaactc	gatcctggcg	tccatccaca	3420
acttcgaccc	tgaggccggc	tgcgacaacc	tgaccttcca	gccctgcagc	gagcgcgccc	3480
tggccaacca	caaggcctat	gtcgactcgt	tccgcaacct	ctacgccatc	aacaagggca	3540
tcgcccaggg	caaggccggt	gccgtcgggc	gctactcgga	ggatgtctac	tacaacggca	3600
acccgtggta	cctggccaac	tttgccgcg	ccgagcagct	ctacgacgcc	atctacgtgt	3660
ggaacaagca	gggtccatc	accgtgacct	cggctctcct	gcccttcttc	cgcgaccttg	3720
tctcgtcgg	cagcaccggc	acctactcca	agagcagctc	gaccttcacc	aacatcgtca	3780
acgccgtcaa	ggcctacgcc	gacggcttca	tcgaggtggc	ggccaagtac	accccgctcca	3840
acggcgcgct	cgccgagcag	tacgaccgca	acacgggcaa	gcccgactcg	gccgcgcgacc	3900
tgacgtggtc	gtactcggcc	ttcctctcgg	ccatcgaccg	ccgcgcgggt	ctcgtccccc	3960
cgagctggcg	ggccagcggt	gccaagagcc	agctgcgcgtc	cacctgctcg	cgcacgcagg	4020
tcgcccggcac	ctacgtcgcc	gccacgagca	cctcgttccc	gtccaagcag	accccggaacc	4080
cctccgcggc	gccctccccg	tccccctacc	cgaccgcctg	cgcggacgct	agcgaggtgt	4140
acgtcacctt	caacgagcgc	gtgtcgaccg	cgtggggcga	gaccatcaag	gtggtgggca	4200
acgtgccggc	gctggggaac	tgggacacgt	ccaaggcgggt	gacctgtcg	gccagcgggt	4260
acaagtcgaa	tgatccctc	tggagcatca	cggtgcccat	caaggcgacg	ggctcggcgg	4320
tgcagtacaa	gtatatcaag	gtcggcacca	acgggaagat	tacttgggag	tcggacccca	4380
acaggagcat	taccctgcag	acggcgctgt	ctgcgggcaa	gtgcgcgcg	cagacggtga	4440
atgattcgtg	gcgttaaaa	ggtgggcgcg	ccgaccacgc	tttcttgtac	aaagtgggtga	4500
tcgcgccagc	tccgtgcgaa	agcctgacgc	accggtagat	tcttggtgag	cccgtatcat	4560
gacggcgcg	ggagctacat	ggccccgggt	gattttatctt	ttttgtatct	acttctgacc	4620
cttttcaaat	atacggtaaa	ctcatctttc	actggagatg	cggcctgctt	ggtattgcga	4680
tggtgtcagc	tgggcaaat	gtggctttcg	aaaacacaaa	acgattcctt	agtagccatg	4740
cattttaaga	taacggaata	gaagaaagag	gaaattaaaa	aaaaaaaaaa	aacaaacatc	4800
ccgttcataa	cccgtagaat	cgccgctctt	cgtgtatccc	agtaccagtt	tattttgaat	4860
agctcgcccc	ctggagagca	tcctgaatgc	aagtaacaac	cgtagaggct	gacacggcag	4920
gtgttgctag	ggagcgtcgt	gttctacaag	gccagacgct	ttcgcggttg	atataatgt	4980
atgtttgact	gcaggctgct	cagcgacgac	agtcaagttc	gccctcgctg	cttgtgcaat	5040
aatcgacgtg	gggaagccac	accgtgactc	ccatctttca	gtaaagctct	gttggtgttt	5100
atcagcaata	cacgtaattt	aaactcgtta	gcattggggt	gatagcttaa	ttaccgttta	5160
ccagtgccat	ggtctgcag	ctttccttgg	cccgtaaaa	tcggcggaagc	cagccaatca	5220
ccagctaggc	accagtaaaa	ccctataatt	agctctcttat	caacaccatc	cgtcccccg	5280
ggatcaatga	ggagaatgag	ggggatgcgg	ggctaaagaa	gcctacataa	ccctcatgcc	5340
aactcccagt	ttacactcgt	cgagccaaca	tcctgactat	aagctaacac	agaatgcctc	5400
aatcctggga	agaactggcc	gctgataagc	gcgcccgcct	cgcaaaaacc	atccctgatg	5460
aatggaaagt	ccagacgctg	cctgcggaag	acagcgttat	tgatttccca	aagaaatcgg	5520
ggatcctttc	agaggccgaa	ctgaagatca	cagaggcctc	cgtgcagat	cttgtgtcca	5580
agctggcggc	cggagagtgt	acctcgggtg	aagttacgct	agcattctgt	aaacgggcag	5640
caatcgccca	gcagttagta	gggtcccttc	tacctctcag	ggagatgtaa	caacgccacc	5700
ttatgggact	atcaagctga	cgtggcttc	tgtgcagaca	aactgcgccc	acgagttctt	5760
ccctgacgcc	gctctcgcg	aggcaaggga	actcgatgaa	tactacgcaa	agcacaagag	5820
acccgttggt	ccactccatg	gcctcccat	ctctctcaaa	gaccagcttc	gagtcaagg	5880
acaccgttgc	ccctaagtgc	ttagatgtcc	ctttttgtca	gctaacatat	gccaccagg	5940
ctacgaaaca	tcaatgggct	acatctcatg	gctaaacaag	tacgacgaag	gggactcggt	6000
tctgacaacc	atgctccgca	aagccgggtg	cgtcttctac	gtcaagacct	ctgtcccgca	6060
gaccctgatg	gtctgcgaga	cagtcaacaa	catcatcggt	cgcaccgtca	accacgcaa	6120
caagaactgg	tcgtgcggcg	gcagttctgt	tggtgaggg	gcgatcgttg	ggattcgcrv	6180
tggtggcgct	atcgggtgag	gaacggatat	cgggtggctg	attcgagtgc	cggcccggtt	6240
caacttcctg	tacgggtctaa	ggccgagtc	tgggcggctg	ccgatgcaa	agatggcgaa	6300
cagcatggag	ggtcaggaga	cggtgcacag	cgtgttcggg	ccgattacgc	actctgttga	6360
gggtgagtcc	ttcgctctct	ccttcttttc	ctgctctata	ccaggcctcc	actgtcctcc	6420
tttcttgctt	tttatactat	atacagagacc	ggcagtcact	gatgaagtat	gttagacctc	6480

GC830-PCT.txt

cgccctcttca	ccaaatccgt	cctcggtcag	gagccatgga	aatacgactc	caaggtcatc	6540
cccattgccct	ggcgccagtc	cgagtcggac	attattgcct	ccaagatcaa	gaacggcggg	6600
ctcaatatcgc	gctactacaa	cttcgacggc	aatgtccttc	cacaccctcc	tatcctgogc	6660
ggcgtggaaa	ccaccgtcgc	cgcactcgcc	aaagccggtc	acaccgtgac	cccgtggacg	6720
ccatacaagc	acgatttcgg	ccacgatctc	atctcccata	tctacgcggc	tgacggcagc	6780
crvgccgacg	taatgcgcga	tatcagtgca	tccggcgagc	cggcgattcc	aaatatcaaa	6840
gacctactga	acccgaacat	caaagctgtt	aacatgaacg	agctctggga	cacgcatctc	6900
cagaagtgga	attaccagat	ggagtacctt	gagaaatggc	gggaggctga	agaaaaggcc	6960
gggaagggaac	tggacgccat	catcgcgccg	attacgccta	ccgctgcggg	acggcatgac	7020
cagttccgggt	actatgggta	tgcctctgtg	atcaacctgc	tggatttcac	gagcgtgggt	7080
gttccgggtta	cctttgcgga	taagaacatc	gataagaaga	atgagagttt	caaggcgggt	7140
agtgcagcttg	atgccttcgt	gcaggaagag	tatgatccgg	aggcgtacca	tggggcaccg	7200
gttgcaattgc	aggttatcgg	acggagactc	agtgaagaga	ggacgttggc	gattgcagag	7260
gaagtgggga	agttgctggg	aaatgtgggtg	actccatagc	taataagtgt	cagatagcaa	7320
tttgacacaag	aaatcaatac	cagcaactgt	aaataagcgc	tgaagtgacc	atgccatgct	7380
acgaaagagc	agaaaaaac	ctgccgtaga	accgaagaga	tatgacacgc	ttccatctct	7440
caaagggaaga	atcccttcag	ggttgcgttt	ccagtctaga	cacgtataac	ggcacaagtg	7500
tctctcacca	aatgggttat	atctcaaattg	tgatctaagg	atggaaagcc	cagaatatcg	7560
atcgcgcgca	gatccatata	tagggccccg	gttataatta	cctcagggtcg	acgtcccctg	7620
gccattcgaa	ttcgtaatca	tggatcatagc	tgtttcctgt	gtgaaattgt	tatccgctca	7680
caattccaca	caacatacga	gccggaagca	taaagtgtaa	agcctggggg	gcctaattgag	7740
tgagctaact	cacattaatt	gcgttgcgct	cactgccgc	tttccagtcg	ggaaacctgt	7800
cgtgccagct	gcattaatga	atcgccaac	gcgcggggag	aggcggtttg	cgtattgggc	7860
gctcttccgc	ttcctcgctc	actgactcgc	tgcgctcggt	cgttcggctg	cggcgagcgg	7920
tatcagctca	ctcaaaggcg	gtaatacggg	tatccacaga	atcaggggat	aacgcaggaa	7980
agaacatgtg	agcaaaaggc	cagcaaaagg	ccaggaaccg	taaaaaggcc	gcgttgctgg	8040
cgtttttcca	taggctccgc	ccccctgacg	agcatcacaa	aaatcgacgc	tcaagtcaga	8100
ggtggcgaaa	cccgacagga	ctataaagat	accaggcggt	tccccctgga	agctccctcg	8160
tgcgctctcc	tgttccgacc	ctgccgctta	ccggataacct	gtccgccttt	ctcccttcgg	8220
gaagcgtggc	gctttctcat	agctcacgct	gtaggtatct	cagttcgggtg	taggtcgctc	8280
gctccaagct	gggtgtgtg	cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	8340
gttaactatcg	tcttgagtcc	aacccggtaa	gacacgactt	atcgccactg	gcagcagcca	8400
ctggtaacag	gatttagcaga	gcgaggtatg	taggcgggtgc	tacagagttc	ttgaagtggg	8460
ggcctaacta	cggctacact	agaagaacag	tatttggtat	ctgcgctctg	ctgaagccag	8520
ttaccttcgg	aaaaagagtt	ggtagctctt	gatccggcaa	acaaaccacc	gctggtagcg	8580
gtggtttttt	tgtttgcaag	cagcagatta	cgcgagaaaa	aaaaggatct	caagaagatc	8640
ctttgatctt	ttctacgggg	tctgacgctc	agtggaacga	aaactcacgt	taagggtatt	8700
tggatcatgag	attatcaaaa	aggatcttca	cctagatcct	tttaaattaa	aaatgaagtt	8760
ttaaatcaat	ctaaagtata	tatgagtaaa	cttgggtctga	cagttaccaaa	tgcttaatca	8820
gtgaggcacc	tatctcagcg	atctgtctat	cttgcttcatc	catagttgcc	tgactccccg	8880
tcgtgtagat	aactacgata	cgggagggct	taccatctgg	ccccagtgct	gcaatgatac	8940
cgcgagaccc	acgctcacgc	gctccagatt	tatcagcaat	aaaccagcca	gccggaaggg	9000
ccgagcgcag	aagtggctct	gcaactttat	ccgcctccat	ccagtctatt	aattgttgcc	9060
gggaagctag	agtaagtagt	tcgccagtta	atagtttgcg	caacgttggt	gccattgcta	9120
caggcatcgt	ggtgtcacgc	tcgtcgtttg	gtatggcttc	attcagctcc	ggttcccaac	9180
gatcaaggcg	agttacatga	tcccccatgt	tgtgcaaaaa	agcggttagc	tccttcgggtc	9240
ctccgatcgt	tgtcagaagt	aagttggccg	cagtggttatc	actcatgggt	atggcagcac	9300
tgcataattc	tcttactgtc	atgccatccg	taagatgctt	ttctgtgact	ggtgagtact	9360
caaccaagtc	attctgagaa	tagtgtatgc	ggcgaccgag	ttgctcttgc	ccggcgtaaa	9420
tacgggataa	taccgcgcca	catagcagaa	ctttaaaagt	gctcatcatt	ggaaaacggt	9480
cttcggggcg	aaaactctca	aggatcttac	cgctgttgag	atccagttcg	atgtaaccca	9540
ctcgtgcacc	caactgatct	tcagcatctt	ttactttcac	cagcgtttct	gggtgagcaa	9600
aaacaggaag	gcaaaatgcc	gcaaaaaagg	gaataagggc	gacacggaaa	tgttgaatac	9660
tcatactctt	cctttttcaa	tattattgaa	gcatttatca	gggttattgt	ctcatgagcg	9720
gatacatatt	tgaatgtatt	tagaaaaata	aacaaatagg	ggttccgcgc	acatttcccc	9780
gaaaagtgcc	acctgacgtc	taagaaacca	ttattatcat	gacattaacc	tataaaaaata	9840
ggcgtatcac	gaggcccttt	cgtctcgcgc	gtttcgggtga	tgacgggtgaa	aacctctgac	9900
acatgcagct	cccgagacgc	gtcacagctt	gtctgtgaagc	ggatgccggg	agcagacaag	9960
cccgtcaggg	cgcgtcagcg	ggtgttggcg	ggtgtcgggg	ctggcttaac	tatgcggcat	10020

GC830-PCT.txt

cagagcagat	tgtactgaga	gtgcaccata	aaattgtaaa	cgttaatatt	ttgttaaaat	10080
tcgcgttaaa	tttttgttaa	atcagctcat	tttttaacca	ataggccgaa	atcggcaaaa	10140
tcccttataa	atcaaaaagaa	tagcccgaga	tagggttgag	tgttggtcca	gtttggaaca	10200
agagtccact	attaaagaac	gtggactcca	acgtcaaagg	gcgaaaaacc	gtctatcagg	10260
gcgatggccc	actacgtgaa	ccatcaccca	aatcaagttt	tttggggtcg	aggtgccgta	10320
aagcactaaa	tcggaaccct	aaagggagcc	cccgatttag	agcttgacgg	ggaaagccgg	10380
cgaacgtggc	gagaaaggaa	gggaagaaag	cgaaaggagc	gggcgctagg	gcgctggcaa	10440
gtgtagcgg	cacgctgcgc	gtaaccacca	caccgcgcgc	gcttaatgcg	ccgctacagg	10500
gcgcgtacta	tggttgcttt	gacgtatgcg	gtgtgaaata	ccgcacagat	gcgtaaggag	10560
aaaataccgc	atcaggcgcc	attcgccatt	caggctgcgc	aactgttggg	aagggcgatc	10620
ggtgcggggc	tcttcgctat	tacgccagct	ggcgaaaggg	ggatgtgctg	caaggcgatt	10680
aagttgggta	acgccagggt	tttcccagtc	acgacgttgt	aaaacgacgg	ccagtgcc	10739

<210> 5

<211> 2169

<212> DNA

<213> *Aspergillus awamori* var. *kawachi*

<400> 5

atgtcgttcc	gatctcttct	cgccctgagc	ggccttgtct	gctcgggggt	ggcaagtgtg	60
atttccaagc	gcgcgacctt	ggattcgtgg	ttgagcaacg	aagcgaccgt	ggcccgtact	120
gcgatcctga	ataacatcgg	ggcgagcgg	gcttggtgtg	cgggcgcgga	ctctggcatt	180
gtcgttgcca	gtcccagcac	cgataaccgc	gactgtatgt	tttgagttcg	gattatgaat	240
gtgtcttgg	tgattgatgc	tgactggcgt	gtcttttgat	gattgtagac	ttctacacct	300
ggactcgcga	ctctgggtctc	gtcatcaaga	ccctcgtcga	cctcttcgcg	aatggagata	360
ctgatctcct	ttccaccatt	gagcactaca	tctcctctca	ggcaattatt	caggggtgtca	420
gtaacccttc	tggatgatctg	tccagcgggtg	gtcttggtga	gcccaggttc	aatgtcgtatg	480
agactgccta	caccggttct	tggggacggc	cgcagcgtga	tggctcctgcc	ctgagagcaa	540
ctgctatgat	cggttttggg	cagtggctgc	ttgtatgttc	tccacctcct	tgcgtctgat	600
ctgcaacata	tgtagccgac	tggtcaggac	aatggctaca	ccagcgtcgc	aacagagatt	660
gtttggcccc	tcgttaggaa	cgacctgtcg	tatgtggctc	agtactggaa	ccagacggga	720
tatggtgtgt	ttgattgatc	ggggttcaag	ggtgtttgtg	catcgagct	aacttcgcgg	780
tcgcagatct	ctgggaagaa	gttaatggct	cgctcttctt	cactattgcc	gtgcaacacc	840
gcgcctcgt	cgaaggtagt	gccttcgcga	cgcccgctcg	ctcgtcctgc	tcttggtgtg	900
attcgcaggc	acctcagatt	ctctgttact	tgcagtcctt	ctggaccggc	agctacatcc	960
tggccaactt	tgacagcagc	cgttccggca	aggacacaaa	cacctcctcg	ggaagcatcc	1020
acacctttga	tcctgaggct	ggatgcgacg	actccacctt	ccagccctgc	tccccgcgtg	1080
cgctcgccaa	ccataaggag	gttgtagact	ctttccgctc	gatctatact	ctcaacgatg	1140
gctcagtgta	cagttaggcg	gttgccggtcg	gtcggtaacc	tgaggatagc	tactacaacg	1200
gcaaccgtg	gttcctgtgc	accttggtcg	ccgcggaaaca	gctgtacgat	gctctgtacc	1260
agtgggacaa	gcaggggtcg	ttggagatca	cagacgtgtc	acttgacttc	ttcaaggctc	1320
tgtacagtgg	tgctgccacc	ggcacgtact	cttcgtccag	ctcgacctat	agcagcattg	1380
tgagtgccgt	caagactttc	gctgatgggt	ttgtttctat	tgtggtaagt	ctacgctaga	1440
cgagcgtcga	tatttacaga	gggtgcgtac	taacaggatt	aggaaactca	cgccgcaagc	1500
aacggctctc	tgtctgagca	attcgacaag	tctgatggcg	acgagctttc	tgctcgcgat	1560
ctgacctgg	cttacgtgc	tctgctgacc	gccaacaacc	gtcgtaattc	tgctcgtgcc	1620
ccgtcttggg	gtgagacctc	tgccagcagc	gtgcccggca	cctgtgcggc	tacctctgcc	1680
tctggtacct	acagcagtg	gaccgtcacc	tctgtggcga	gcacgtggc	tactggtggc	1740
accactacga	cggctactac	cactggatcg	ggcggcgtga	cctcgaccag	caagaccacc	1800
acaactgcta	gtaagaccag	caccactacg	tctcgcacct	cctgcaccac	ccccactgcc	1860
gtagctgtga	cctttgatct	gacggcgacc	accacctacg	gcgagaacat	ctacctggtc	1920
gggtcgatct	ctcagctcgg	tgactgggag	accagcgatg	gcatagctct	gagcgtgac	1980
aagtacactt	ccagcaaccc	gctttggtat	gtaactgtga	ctctgccggc	tggtagtca	2040
tttgagtaca	agttcatccg	cgtcgagagc	gatgactccg	tggagtgagg	gagcgacccg	2100
aaccgggaat	acaccgttcc	tcaggcgtgc	ggcgagtcga	ccgcgacgg	gaccgacacc	2160
tggcggtag						2169

<210> 6

<211> 24

GC830-PCT.txt

<212> PRT

<213> Aspergillus awamori var. kawachi

<400> 6

Met Ser Phe Arg Ser Leu Leu Ala Leu Ser Gly Leu Val Cys Ser Gly
 1 5 10 15
 Leu Ala Ser Val Ile Ser Lys Arg
 20

<210> 7

<211> 615

<212> PRT

<213> Aspergillus awamori var. kawachi

<400> 7

Ala Thr Leu Asp Ser Trp Leu Ser Asn Glu Ala Thr Val Ala Arg Thr
 1 5 10 15
 Ala Ile Leu Asn Asn Ile Gly Ala Asp Gly Ala Trp Val Ser Gly Ala
 20 25 30
 Asp Ser Gly Ile Val Val Ala Ser Pro Ser Thr Asp Asn Pro Asp Tyr
 35 40 45
 Phe Tyr Thr Trp Thr Arg Asp Ser Gly Leu Val Ile Lys Thr Leu Val
 50 55 60
 Asp Leu Phe Arg Asn Gly Asp Thr Asp Leu Leu Ser Thr Ile Glu His
 65 70 75 80
 Tyr Ile Ser Ser Gln Ala Ile Ile Gln Gly Val Ser Asn Pro Ser Gly
 85 90 95
 Asp Leu Ser Ser Gly Gly Leu Gly Glu Pro Lys Phe Asn Val Asp Glu
 100 105 110
 Thr Ala Tyr Thr Gly Ser Trp Gly Arg Pro Gln Arg Asp Gly Pro Ala
 115 120 125
 Leu Arg Ala Thr Ala Met Ile Gly Phe Gly Gln Trp Leu Leu Asp Asn
 130 135 140
 Gly Tyr Thr Ser Ala Ala Thr Glu Ile Val Trp Pro Leu Val Arg Asn
 145 150 155 160
 Asp Leu Ser Tyr Val Ala Gln Tyr Trp Asn Gln Thr Gly Tyr Asp Leu
 165 170 175
 Trp Glu Glu Val Asn Gly Ser Ser Phe Thr Ile Ala Val Gln His
 180 185 190
 Arg Ala Leu Val Glu Gly Ser Ala Phe Ala Thr Ala Val Gly Ser Ser
 195 200 205
 Cys Ser Trp Cys Asp Ser Gln Ala Pro Gln Ile Leu Cys Tyr Leu Gln
 210 215 220
 Ser Phe Trp Thr Gly Ser Tyr Ile Leu Ala Asn Phe Asp Ser Ser Arg
 225 230 235 240
 Ser Gly Lys Asp Thr Asn Thr Leu Leu Gly Ser Ile His Thr Phe Asp
 245 250 255
 Pro Glu Ala Gly Cys Asp Asp Ser Thr Phe Gln Pro Cys Ser Pro Arg
 260 265 270
 Ala Leu Ala Asn His Lys Glu Val Val Asp Ser Phe Arg Ser Ile Tyr
 275 280 285
 Thr Leu Asn Asp Gly Leu Ser Asp Ser Glu Ala Val Ala Val Gly Arg
 290 295 300
 Tyr Pro Glu Asp Ser Tyr Tyr Asn Gly Asn Pro Trp Phe Leu Cys Thr
 305 310 315 320
 Leu Ala Ala Ala Glu Gln Leu Tyr Asp Ala Leu Tyr Gln Trp Asp Lys
 325 330 335
 Gln Gly Ser Leu Glu Ile Thr Asp Val Ser Leu Asp Phe Phe Lys Ala
 340 345 350

GC830-PCT.txt

Leu Tyr Ser Gly Ala Ala Thr Gly Thr Tyr Ser Ser Ser Ser Thr
 355 360 365
 Tyr Ser Ser Ile Val Ser Ala Val Lys Thr Phe Ala Asp Gly Phe Val
 370 375 380
 Ser Ile Val Glu Thr His Ala Ala Ser Asn Gly Ser Leu Ser Glu Gln
 385 390 395 400
 Phe Asp Lys Ser Asp Gly Asp Glu Leu Ser Ala Arg Asp Leu Thr Trp
 405 410 415
 Ser Tyr Ala Ala Leu Leu Thr Ala Asn Asn Arg Arg Asn Ser Val Val
 420 425 430
 Pro Pro Ser Trp Gly Glu Thr Ser Ala Ser Ser Val Pro Gly Thr Cys
 435 440 445
 Ala Ala Thr Ser Ala Ser Gly Thr Tyr Ser Ser Val Thr Val Thr Ser
 450 455 460
 Trp Pro Ser Ile Val Ala Thr Gly Gly Thr Thr Thr Thr Ala Thr Thr
 465 470 475 480
 Thr Gly Ser Gly Gly Val Thr Ser Thr Ser Lys Thr Thr Thr Thr Ala
 485 490 495
 Ser Lys Thr Ser Thr Thr Thr Ser Ser Thr Ser Cys Thr Thr Pro Thr
 500 505 510
 Ala Val Ala Val Thr Phe Asp Leu Thr Ala Thr Thr Thr Tyr Gly Glu
 515 520 525
 Asn Ile Tyr Leu Val Gly Ser Ile Ser Gln Leu Gly Asp Trp Glu Thr
 530 535 540
 Ser Asp Gly Ile Ala Leu Ser Ala Asp Lys Tyr Thr Ser Ser Asn Pro
 545 550 555 560
 Leu Trp Tyr Val Thr Val Thr Leu Pro Ala Gly Glu Ser Phe Glu Tyr
 565 570 575
 Lys Phe Ile Arg Val Glu Ser Asp Asp Ser Val Glu Trp Glu Ser Asp
 580 585 590
 Pro Asn Arg Glu Tyr Thr Val Pro Gln Ala Cys Gly Glu Ser Thr Ala
 595 600 605
 Thr Val Thr Asp Thr Trp Arg
 610 615

<210> 8
 <211> 28
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> primer

<400> 8
 caacatgcat accttctcca agctcctc

28

<210> 9
 <211> 25
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> primer

<400> 9
 ttaacgccac gaatcattca ccgtc

25

<210> 10
 <211> 25

GC830-PCT.txt

<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 10
caccatgtcg ttccgatctc ttctc

25

<210> 11
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 11
ctaccgccag gtgtcggta c

21